

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Факультет електроніки

Кафедра акустики та акустoeлектроніки

«На правах рукопису»

УДК 621.391.83

«До захисту допущено»

Завідувач кафедри

_____ Дідковський В.С.

(підпис)

(ініціали, прізвище)

“ _____ ” _____ 2018 р.

Магістерська дисертація

зі спеціальності (спеціалізації) 171 Електроніка _____

на тему: Об’єктивне та суб’єктивне оцінювання якості музичних сигналів
обмежених смугою частот _____

Виконав : студент 6 курсу, групи ДГ-62м

_____ Юраков Микита Павлович _____

(підпис)

Науковий керівник проф. каф. А та АЕ, д.т.н., Продеус А.М. _____

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Консультант _____

(назва розділу)

_____ (науковий ступінь, вчене звання, , прізвище, ініціали)

(підпис)

Рецензент проф. КММТ НАУ, к.т.н., доцент, Мелешко М.А. _____

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних
посилань.

Студент _____

(підпис)

Київ – 2018 року

**Національний технічний університет України
«Київський політехнічний інститут
імені Ігоря Сікорського»**

Інститут/факультет _____ електроніки _____
(повна назва)

Кафедра _____ акустики та акустoeлектроніки _____
(повна назва)

Рівень вищої освіти – другий (магістерський) за освітньо-професійною
(освітньо-науковою) програмою

Спеціальність (спеціалізація) _____ 171 Електроніка _____
(код і назва)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Дідковський В.С.
(підпис) (ініціали, прізвище)
«__» _____ 2018 р.

ЗАВДАННЯ

на магістерську дисертацію студенту

_____ Юракову Микиті Павловичу _____
(прізвище, ім'я, по батькові)

1. Тема дисертації Об'єктивне та суб'єктивне оцінювання якості
музичних сигналів обмежених смугою частот
науковий керівник дисертації Продеус А.М., д.т.н., проф. кафедри А та АЕ,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «09» лютого 2018 р. № 691-с

2. Строк подання студентом дисертації «11» травня 2018 року

3. Об'єкт дослідження Алгоритми об'єктивного оцінювання якості
музичних сигналів

4. Предмет дослідження (Вихідні дані – для магістерської дисертації за
освітньо-професійною програмою) _____
Оцінка якості сигналів отриманих різними алгоритмами, їх порівняння та
співставлення.

5. Перелік завдань, які потрібно розробити _____
 1) Аналітичний огляд літературних джерел _____
 2) Розробка плану проведення експериментального дослідження _____
 3) Підготовка до проведення експериментального дослідження _____
 4) Проведення дослідження та аналіз отриманих даних _____
 5) Оформлення результатів дослідження у вигляді публікацій та _____
 магістерської дисертації _____

6. Перелік графічного (ілюстративного) матеріалу _____ презентація _____

7. Орієнтовний перелік публікацій _____
 Юраков М.П., Об'єктивне та суб'єктивне оцінювання якості _____
 спотворених музичних сигналів // XI Міжнародна науково – технічна _____
 конференція молодих вчених «Електроніка-2018» – К.: «КПІ імені Ігоря _____
 Сікорського», 2018, - сс.23-26. _____

8. Консультанти розділів дисертації*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

9. Дата видачі завдання _____ 12 лютого 2018 року _____

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	Аналітичний огляд літератури	12.02.2018-25.02.2018	
2	Розробка схеми експериментального дослідження	26.02.2018-11.03.2018	
3	Проведення експериментального дослідження	12.03.2018-15.04.2018	
4	Розроблення стартап-проекту	16.04.2018-29.04.2018	
5	Оформлення ПЗ та презентації	12.04.2018-06.05.2018	

Студент

_____ (підпис)

Юраков М.П.

_____ (ініціали, прізвище)

Науковий керівник дисертації

_____ (підпис)

Продеус А.М.

_____ (ініціали, прізвище)

* Консультантом не може бути зазначено наукового керівника

РЕФЕРАТ

Об'єктивне та суб'єктивне оцінювання якості музичних сигналів обмежених смугою частот // Дисертація на здобуття ступеня вищої освіти «магістра» Юракова М. П. Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», факультет електроніки, кафедра акустики та акустoeлектроніки, група ДГ-62м. – К.:НТУУ «КПІ імені Ігоря Сікорського», 2018. с. – 56, рис. – 12, табл. – 27.

Складність сучасних ліній зв'язку зростає, при цьому зокрема розробляються стандарти, які дозволяють передавати не тільки мову а і музику. Оскільки постачальників послуг в першу чергу цікавить якість сприйнята кінцевим користувачем, то для оцінювання цієї якості використовують різноманітні методи кількісної оцінки якості сигналів.

Метою даної роботи було визначення залежності якості музичних сигналів від ширини смуги частот та знаходження граничного значення смуги частот після якої людина не помітить покращення якості сигналу. А також знаходження відповідності суб'єктивних і об'єктивних показників якості для музичних сигналів.

Для оцінювання якості сигналів обрано один метод суб'єктивного тестування, іменований «оцінкою ступеня деградації» (Degradation Category Rating - DCR) та чотири показники, такі як: сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio), логарифмічне спектральне спотворення LSD (Logarithmic Spectral Distortion), барк-спектральне спотворення BSD (Bark Spectral Distortion) і перцептивна оцінка якості звуку PEAQ (Perceptual Evaluation Audio Quality).

В роботі шляхом експериментальних досліджень отримані суб'єктивні та об'єктивні оцінки якості музичних сигналів в залежності від смуги пропускання для достатньо широкого набору тестових сигналів, а також проведено зіставлення результатів суб'єктивних та об'єктивних показників. В результаті оцінки коефіцієнта кореляції та побудови карт відповідності між

суб'єктивними та об'єктивними оцінками показано перевагу у використанні карт відповідності замість коефіцієнта кореляції.

Результати роботи були оприлюднені на XI міжнародна науково-технічній конференції молодих вчених, що проходила 3-5 квітня 2018 року та опубліковані в збірнику матеріалів конференції, у статті «Об'єктивне та суб'єктивне оцінювання якості спотворених музичних сигналів»

Ключові слова: оцінювання якості, телефонія, зв'язок, музика, кодек, degradation category rating, segmental signal to noise ratio, logarithmic spectral distortion, bark spectral distortion, perceptual evaluation audio quality.

ABSTRACT

Quality evaluation of music signals restricted by frequency band // Thesis for a master's degree by Yurakov N. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Faculty of Electronics, Department of acoustics and acoustoelectronics, group DG-62m. - K.: NTUU "Igor Sikorsky Kyiv Polytechnic Institute", 2018. p. - 56, fig. - 12, tab. - 27.

The complexity of modern communication lines grows, in particular, developing a standards that allows the transfer not only a speech but music. Service providers are primarily interested in the quality of the end-user's perception, various methods for quantifying signal quality are used to evaluate this quality.

The purpose of this work was to determine the quality dependence of musical signals on the frequency band and the finding of the limit value of the bandwidth after which a person does not notice improvement in signal quality. And also finding matching of subjective and objective indicators of quality for musical signals.

For evaluation of signals quality were choosen one method of subjective testing, called the Degradation Category Rating (DCR), and four methods of ojective testing, such as: Segmental Signal to Noise Ratio Segmental Signal to Noise Ratio, logarithmic spectral distortion LSD, is selected to evaluate signal quality. (Logarithmic Spectral Distortion), Bark Spectral Distortion (BSD), and perceptual evaluation of sound quality (PEAQ) (Perceptual Evaluation Audio Quality).

In the work, through experimental research subjective and objective estimates of the quality of musical signals, depending on the bandwidth for a sufficiently large set of test signals, were obtained, as well as a comparison of the results of subjective and objective indicators. As a result of the estimation of the correlation coefficient and the mapping of correspondence between subjective and objective estimates, it is shown that the advantage of using matching maps instead of the correlation coefficient is shown.

The results of the work were announced at the XI International Scientific and Technical Conference of Young Scientists, which took place on 3-5 April 2018 and

published in the collection of conference materials, in the article "Objective and subjective evaluation of the quality of distorted musical signals"

Key words: quality evaluation, communication, telephony, music, codec, degradation category rating, segmental signal to noise ratio, logarithmic spectral distortion, bark spectral distortion, perceptual evaluation audio quality.

ЗМІСТ

ПЕРЕЛІК СКОРЧЕНЬ	10
ВСТУП	11
РОЗДІЛ 1. МЕТОДИ ОЦІНЮВАННЯ ЯКОСТІ СИГНАЛІВ	19
1.1. Суб'єктивні методи оцінювання якості сигналів	20
1.1.1. Absolute Category Rating (ACR).....	20
1.1.2. Degradation Category Rating - DCR.....	21
1.1.3. Оцінювання малих спотворень.....	21
1.1.4. MUSHRA.....	22
1.2. Об'єктивні методи оцінювання якості сигналів	22
1.2.1. Segmental Signal to Noise Ratio - SSNR.....	23
1.2.2. Logarithmic Spectral Distortion - LSD	23
1.2.3. Bark Spectral Distortion - BSD	24
1.2.4. Perceptual Evaluation of Audio Quality - PEAQ	25
1.3. Висновки	33
РОЗДІЛ 2. ОРГАНІЗАЦІЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ.....	35
2.1. Підготовчий етап.....	35
2.2. Проведення експерименту	38
2.3. Висновки	39
РОЗДІЛ 3. АНАЛІЗ РЕЗУЛЬТАТІВ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ.....	40
3.1. Результати суб'єктивного та об'єктивного оцінювання	40
3.2. Співставлення результатів суб'єктивного та об'єктивного оцінювання	43
3.3. Висновок	46

РОЗДІЛ 4. РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ	47
4.1 Опис ідеї проекту	47
4.2 Технологічний аудит ідеї проекту	47
4.3 Аналіз ринкових можливостей запуску стартап-проекту	47
4.4. Розроблення ринкової стратегії проекту	49
4.5. Розроблення маркетингової програми стартап-проекту	50
4.6 Висновки	51
ВИСНОВОК.....	52
СПИСОК ПОСИЛАНЬ	54

ПЕРЕЛІК СКОРЧЕНЬ

АЧХ – амплітудно-частотна характеристика

MCE – Міжнародний союз електрозв'язку

ІКМ – Імпульсно-Кодова Модуляція

CD – Компакт диск

VoIP – Voice over Internet Protocol

ARC – Absolute Category Rating

DCR – Degradation Category Rating

MUSHRA – MUltiple Stimuli with Hidden Reference and Anchor

SPL – рівень звукового тиску

SSNR – Segmental Signal to Noise Ratio

LSD – Log-Spectral Distortion

BSD –Bark Spectral Distortion

PEAQ –Perceptual Evaluation of Audio Quality

ODG –Overall Difference Grade

ВСТУП

Протягом останніх чотирьох десятиліть системи передачі мови значно змінились, у літературі ця трансформація називається "цифровою революцією". Першим головним поштовхом стало введення цифрової передачі даних. Наприклад, широкосмугова цифрова мережа з інтеграцією послуг (Integrated Services Digital Network - ISDN) була введена для заміни аналогової передачі на стаціонарній телефонній мережі [1]. Далі у всьому світі поширились системи мобільного зв'язку, такі як глобальна система мобільного зв'язку (GSM). І мобільна, і стаціонарна телефонії побудовані на мережах з комутацією каналів. Двоє співрозмовників пов'язані фізичною схемою, яка тримає весь дзвінок. В додаток до стаціонарних та мобільних мережі, були введені комп'ютерні мережі, такі як Інтернет. Ці комп'ютерні мережі, також відомі як мережі з комутацією пакетів, засновані на переривчастій передачі пакетів даних[1],[2]. Провайдери адаптували мережі з комутацією пакетів для передачі мови, що використовують Voice over Internet Protocol (VoIP) для передачі. В недалекому майбутньому, VoIP-передачі стануть основною службою мовлення серед усіх видів передачі мовлення. Незважаючи на важливі досягнення в технологіях пов'язаних з телефонією протягом останніх чотирьох десятиліть, найімовірніше знатиме наступне десятиліття збільшення послуг зв'язку, включаючи нові можливості.

Однак ці три мовні послуги мають дуже різну природу. Мережа з комутацією каналів має фіксовану структуру і запровадження нових систем обробки навряд чи можливо. Сприйняття якості користувачами стаціонарних комутаційних мереж було вивчено в минулому. Ці дослідження спираються на результати слухових експериментів і були спрямовані на детальне визначення різноманітних деградацій, запроваджених такими мережами. [2]

Ці деградації є постійними протягом усієї розмови. Мобільна і VoIP-мережі мають гнучкість щодо впровадження нової системи обробки мови, наприклад, VoIP-мережі дозволяють передавати мову широкосмуговою

лінією. Передача широкосмуговою лінією є значним поліпшенням у порівнянні з фіксованою мережею з комутацією каналів. Крім того, мережа GSM забезпечує більшу мобільність для користувача.

З точки зору користувача, ці нові мережі, вносять в основному лише переваги. Однак ці мережі вводять нові типи деградації сигналів, що передаються. Наприклад, голос диктора може обірватися, коли пакети втрачаються під час передачі. До того ж, інформація, що стосується навколишнього оточення диктора є більш важливою у мобільній мережі, ніж у стаціонарній мережі. Зв'язок крізь різні мережі може навіть поєднувати деградації цих мереж.

Оцінка якості мовлення охоплює різні галузі навчання, такі як психоакустика, обробка сигналів, тощо. У телекомунікаційній галузі якість мовлення в основному використовується як метод оцінки продуктів, що використовують мовні технології. Як сприймається якість мови, переданої телефонною мережею вивчається протягом останніх десятиліть [1], [2]. Визначені основні фактори, що впливають на якість мови. В результаті, якість мовлення має багатовимірний характер. За ці роки розроблена велика кількість перцептивних функцій, за яким можна порівнювати всі мережі. Однак з цією метою були проведені численні слухові випробування. Ці випробування, на жаль, вимагають багато часу та витрат. Однак є велика необхідність постачальників телекомунікаційних послуг вимірювати якість передачі мови системами. Такі методи вимірювання можуть бути корисними у двох різних випадках:

- 1) під час планування та розробки нових систем обробки мови (наприклад, мовні кодеки);
- 2) для моніторингу користування телефонними мережами.

Цими технологіями зацікавлені розробниками систем обробки голосу та засобів зв'язку, а також провайдери мовлення.

Зазвичай інформація про якість системи передачі забезпечується простим інструментальним способом вимірювані параметрів, таких як смуга

пропускання, відношення сигнал / шум або відсоток втрачених пакетів. Проте головний інтерес в основному являє собою якість, що сприймається споживачами. Симуляція сприйняття людиною дуже важке завдання. Для симулювання були розроблені складні методи вимірювання суб'єктивної складової людини в тестах прослуховуваннях, наприклад PSQM або PESQ, описані в ITU-T Rec. P.862 [3]. Вони аналізують мовний сигнал з використанням моделі периферичної слухової системи людини. Такі методи були розроблені для комутованих мереж і мереж з комутацією пакетів. Однак вони показують певні обмеження при порівнянні різних мереж. Методи, що використовуються нині, такі як PESQ, не пристосовані до нових методів, що використовуються в системах обробки мови. Тому є потреба в розвитку загального інструментального методу вимірювання, який може оцінити кожен вживаний тип мережі та системи обробки мови, включаючи їх можливі комбінації.

Як можна побачити якість мови є предметом обговорення з появи перших телефонних ліній і по нині. Однак, передача аудіо-інформації не обмежується однією лише трансляцією мовних програм та телефонією.

Що стосується трансляції музичних передач, то до недавнього часу єдиними системами що використовувались в цьому напрямку були аналогове телебачення та радіомовлення. Однак не можливо заперечити той факт, що використання та популярність цифрового аудіо-мовлення, систем та програм веб-трансляцій зростають неухильно. Через еволюцію цифрового радіо та систем зв'язку, таких як цифрове звукове мовлення (Digital Audio Broadcasting - DAB), (Digital Radio Mondiale- DRM) та наземне цифрове мультимедійне мовлення (T-DMB)[3]. Наразі вище згадані системи радіомовлення замінюють звичайні аналогові радіосистеми у всьому світі.

Алгоритми аудіо кодування або алгоритми стиснення звуку використовуються для отримання компактних цифрових представлень високоякісних (широкосмугових) аудіосигналів для ефективної передачі або зберігання. Центральною метою кодування звуку є представлення сигналу з

мінімальним числом бітів при досягненні високоякісного відтворення сигналу, тобто створення вихідного аудіосигналу, який не можна відрізнити від вхідного, навіть підготовленим слухачем. Введення компакт-диска (CD) на початку 80-х років показало всі переваги цифрового аудіо представлення, включаючи безпрецедентну високу точність, динамічний діапазон і надійність. Ці переваги, однак, прийшли за рахунок високих об'ємів даних. Звичайні системи для компакт-дисків та цифрових аудіокaset (DAT), як правило, кодуються як правило з частотами дискретизації 44,1 або 48 кГц за допомогою імпульсно - кодової модуляції (ІКМ) з роздільною здатністю 16 біт. Це призводить для некомпресованих сигналів до значних вимог щодо швидкості передачі даних 705,6 / 768 кілобіт на секунду (кбіт / с) для монофонічного каналу або 1,41 / 1,54 мегабіт на секунду (Мбіт / с) для стереопар на частотах 44,1 / 48 кГц, відповідно[4].

На жаль, мультимедійні додатки другого покоління і, зокрема, бездротові системи часто піддаються обмеженню пропускної здатності та витрат, що несумісні з високим бітрейтом. Проте, завдяки успіху, отриманому першим поколінням, кінцеві користувачі приходять до очікування звуку якості "CD" відтворення з будь-якої цифрової системи. Тому нові мережеві та бездротові мультимедійні цифрові аудіосистеми повинні зменшуватися кількість даних, що передаються без погіршення якості відтворення. Ці та інші міркування зумовили значні дослідження протягом останніх чотирьох десятиліть до розробки кодеків стиснення, здатних одночасно задовольнити суперечливі вимоги високих коефіцієнтів стиснення та високої якості відтворення аудіосигналів.

Цифрові аудіо системи потокового мовлення та трансляції в Інтернеті додатки розроблені таким чином, щоб бути дуже стійкими до помилок які могли б бути введені по каналу передачі, завдяки використанню високоефективних схем кодування каналів (системи мовлення) або дуже надійний метод передачі на основі протоколу TCP. В таких системах аудіо кодек є основним джерелом деградації сигналів. Найбільш застосованими

схемами кодування з втратами є MP2, HE-AAC, MP3, Ogg Vorbis, AAC-LC і Opus. Бітрейт цих кодеків лежить в межах від 24 до 320 кбіт / с. [5],[6],[7]

В перших версіях систем стиснення з втратами якість звуку, що сприймалася кінцевим користувачем, безпосередньо залежала від бітрейта. Що змушувало телерадіомовлення та веб-програми користуватись найнижчою показниками бітрейту. Та методи кодування з втратами значно вдосконалились з розвитком цифрового мовлення, і тепер підвищення якості з ростом бітрейту вже не є настільки очевидним. Тому вирішальне значення для постачальників телекомунікаційних та мультимедійних послуг є вимірювання та контроль якості, що сприймається кінцевим користувачем в системи цифрового аудіо-мовлення та веб-кастингу. Оскільки користувачі переходять з аналогових систем на ці цифрові, то вони оцікуватимуть більш високої якості, ніж в аналогових системах. Більш детально залежність якості музичних сингалів від бітрейту розглянуто в роботах [8], [9], [10], [11], та [12].

Для зменшення затримок в каналі зв'язку використовують різні методи стиснення даних, що призводить до зменшення якості сигналів. Найбільш ефективним, поширеним і простим методом стиснення є зменшення ширини смуги частот.

Всі ми знаємо що діапазон людського слуху лежить в межах 20-20000 Гц, тому в першу чергу зменшення смуги пропускання спирається на частотну смугу сигналу що передається. Наприклад мовленеві сигнали, зазвичай лежать в діапазоні 100-7000 Гц. Ця смуга пропускання сильно залежить від вироблених звуків. Більша частина енергії звуків сконцентрована у перших трьох формантах (тобто в діапазоні 300-3000 Гц) [1], що відображалось втому, що до введення стандарту мобільного зв'язку 3G та появи VoIP смуга акустичного тракту була обмежена 3-4 кГц. Тоді як шумні приголосні, такі як «ж», «ч», «ш», тощо, мають основну енергії в діапазоні 3 -10 кГц, що і отримало своє відображення в стандарті 3G[12]. Однак, крім вищезгаданих атрибутів, складовою мови є деякі зрозумілі звуки та інші показники, таких як

свистки, клацання, шипіння, тощо. Як можна побачити з рис.1 якість каналів зв'язку має ядро виражену тенденцію до зростання.

В Україні на даний час діє стандарт 3G, що забезпечує смугу пропускання в 7 кГц, а стандарт 4G стане доступним протягом року, та чи достатньо смуги в 14 кГц, яку забезпечує стандарт 4G, для якісної передачі музичних сигналів.

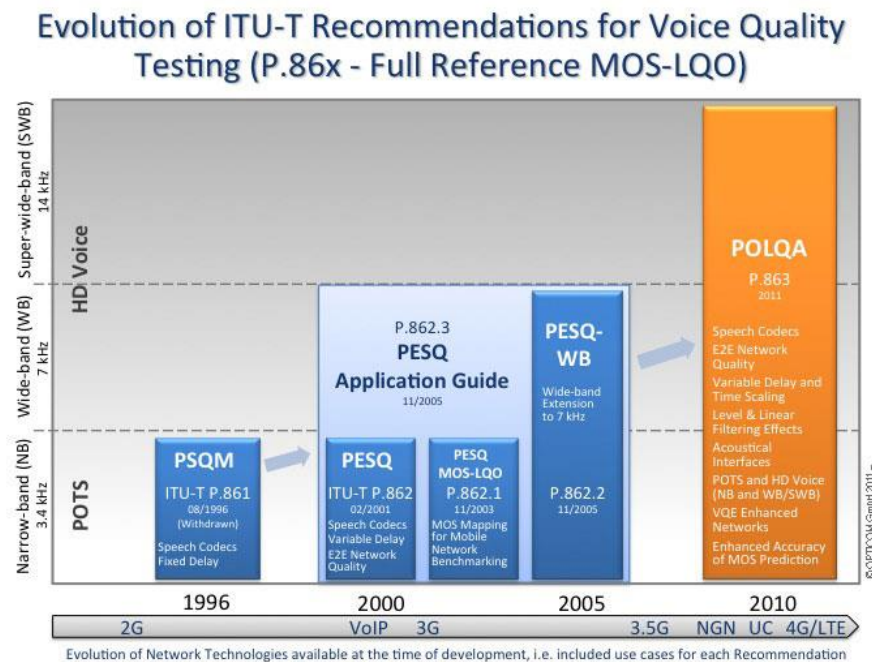


Рис.1. Розвиток рекомендацій МСЕ що до якості ліній мобільного зв'язку та методів оцінювання якості [5]

В цифровому телебаченні та IP-телефонії на даний час звуковий тракт обмежений смугою в 12 кГц, про що йдеться в [13], [14] та [15]. Кодек Opus може підтримувати частоту дискретизації і в 48 кГц, однак не має підтвердження його використання для передачі даних в реальному часі немає. Використання підвищеної частоти дискретизації пов'язане зі збільшенням кількості інформації яку потрібно передати, що може призвести до значних затримок в каналі зв'язку.

Таблиця 1. Смуги пропускання деяких популярних каналів зв'язку[8 -12]

	Канал зв'язку (кодек)		
	DVB-T2 (AAC-LC)	Skype (SILK)	Viber (Opus)
Частота дискретизації	8-24кГц	8-24кГц	8-24кГц
Смуга пропускання	4-12кГц	4-12кГц	4-12кГц

Розширення смуги частот до 20 кГц обумовлено бажанням передавати музику по каналах зв'язку, що поліпшить враження від перегляду музичних передач, але й також важливо при навчанні музикантів на великих відстанях (Австралія-Америка, тощо), адже, як було вище згадано, діапазон людського слуху обмежений смугою 20 – 20000 Гц.

З іншого боку слід зауважити, що ці данні говорять лише про фізичні обмеження нашого слуху і не враховують, звуження частотного діапазону з віком та інших особливостей нашого слуху: маскування, адаптація гучності тощо, які вносять більший вклад в сприйняття якості сигналів. В роботі Бранденбурга [13] даються пояснення кодеків MP3 та AAC та психоакустичних ефектів які враховуються при стисненні аудіо-сигналу, та найцікавішим є припущення про те, що пересічна людина не може помітити різницю між двома складними музичними сигналами один яких має інформацію аж до 20 кГц, а інший обмежений смугою 16 кГц. Проте автор статті наголошує на тому, що наукового підтвердження даній гіпотезі наразі немає.

Також є робота [14] (І. Котвицкий, А. Оса) де автори мали за мету дослідження відповідності суб'єктивних і об'єктивних показників якості для музичних сигналів. Однак в роботі, нажаль, не має фактичного зіставлення цих показників, що залишає прогалину у даному питанні. Також спірним є підбір музичного матеріалу, що піддавався тестуванню, точніше його низькі

представницькість та різноманіття. Окремої уваги потребує і той факт, що всі обрані об'єктивні методи оцінювання працюють тільки в монофонічному режимі, що детально буде розглянуто в розділі 2, однак автори не повідомляють чи враховували вони дану особливість. Натомість, з їх роботи можна отримати орієнтовне значення граничного значення смуги частот для якісного відтворення музики в 14-15 кГц, що й потрібно буде перевірити.

Метою даної роботи є заповнення цих прогалин, тобто визначення залежності якості музичних сигналів від ширини смуги частот та знаходження граничного значення смуги частот після якої людина не помітить покращення якості сигналу. А також знаходження відповідності суб'єктивних і об'єктивних показників якості для музичних сигналів.

РОЗДІЛ 1. МЕТОДИ ОЦІНЮВАННЯ ЯКОСТІ СИГНАЛІВ

Для кількісного визначення сприйнятої людиною якості звуку телекомунікаційного тракту або аудіо кодека, що використовується лінією зв'язку, протягом останніх десятиліть був розроблений ряд методів оцінювання якості сигналів.

Говорячи про оцінювання якості сигналів, зазвичай мають на увазі, що проводиться явне або неявне порівняння спотвореного мовного сигналу з неспотвореним сигналом. Чим більше така відмінність, тим нижче якість.

Більшість методів розроблених для оцінки якості телекомунікаційних каналів досить не нові і в першу чергу передбачали використання мовних сигналів. Більшість методів мають обмеження по ширині полоси пропускання

Прийнято розрізняти [1] дві групи показників якості мови: суб'єктивні (аудіювання) і об'єктивні (інструментальні) (рис.1.1).

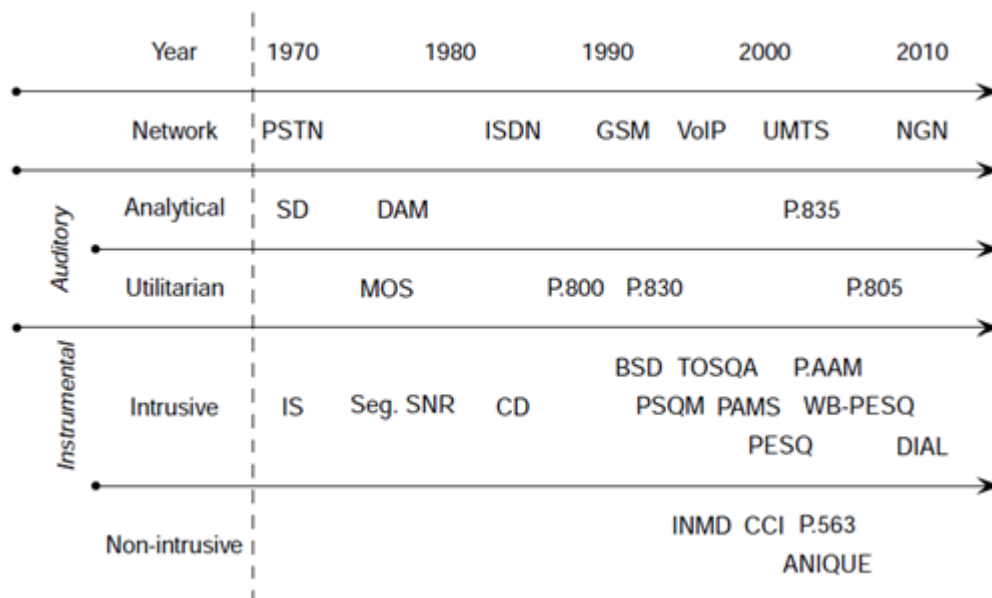


Рис. 1.1. Розвиток методів оцінювання якості сигналів [1]

Суб'єктивні методи поділяються на утилітарні та аналітичні. У методах утилітарного тестування суб'єкти оцінюють інтегральну якість каналу передачі, використовуючи єдину шкалу оцінок. Це дозволяє зіставляти різні

умови обробки. У методах аналітичного тестування спочатку ідентифікуються, а потім кількісно оцінюються перцептивні характеристики інтегральної якості мови та музики. При цьому може використовуватися одна або кілька шкал оцінювання.

Перевагою суб'єктивних показників є те, що оцінка якості мовного сигналу проводиться з використанням слухової системи людини, а недоліками є трудомісткість і дорожня процедура оцінювання. У зв'язку з цим величезне значення набуває використання відносно недорогих об'єктивних (інструментальних) показників, оцінювання яких проводиться двома способами:

- з використанням еталонного сигналу (інтрузивні показники);
- без використання еталонного сигналу (неінтрузивні показники).

1.1. Суб'єктивні методи оцінювання якості сигналів

Як вже згадувалось вище суб'єктивні методи поділяються на утилітарні та аналітичні. Зазвичай перші використовуються при оцінюванні якості кінцевим споживачем, другі ж вимагають залучення експертів, що мають певну підготовку в прослуховуваннях. Виходячи з цього далі будуть розглянуті лише утилітарні методи оцінювання якості сигналів, оскільки важливим для даної роботи є отримання результату, який був би справедливим для більшості людей (непідготованих слухачів).

1.1.1. Absolute Category Rating (ACR)

Тестування систем зв'язку найчастіше здійснюють методом оцінювання за абсолютною шкалою (Absolute Category Rating - ACR), згідно з яким якість прослуховується сигналу оцінюється тільки за результатами прослуховування тестового сигналу - порівняння з «зразковим ідеальним» сигналом при цьому не проводиться. Відповідну п'ятибальну шкалу оцінювання (табл. 2.2) називають MOS-Listening Quality Score (MOS_{LQS}) [18].

Таблиця 1.1. Шкала MOS_{LQS} методу ACR [1]

Quality of the speech	Score
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

1.1.2. Degradation Category Rating - DCR

Навпаки, при оцінюванні методом «оцінка ступеня деградації» (Degradation Category Rating - DCR) прослуховується сигнал порівнюють з еталонним сигналом високої якості - відповідну шкалу (табл. 2.3) називають Degradation Mean Opinion Score (DMOS) [18].

Таблиця 1.2. Шкала методу DMOS [1]

Score	The degradation is ...
5	inaudible
4	audible but not annoying
3	slightly annoying
2	annoying
1	very annoying

1.1.3. Оцінювання малих спотворень

При оцінюванні якості широкосмугових і надширокосмугових сигналів застосовують метод оцінювання малих спотворень, суть якого в тому, що слухачеві пред'являють три звукових зразка А, В і Х, після чого просять сказати, який із зразків, А або В, ближче до зразка Х. [18]

Методи тестування широкосмугових (Wide Band - WB) та надширокосмугових (Super-Wide Band - S-WB) каналів викладені в документі ITU-T Rec. P.830 [19], а також в документах ITU-R Rec. BS.1116-1, 1997. [20]; ITU-R Rec. BS.1284-1, 2003 [21]. При цьому умови тестування відрізняються

від таких для вузькосмугового зв'язку більш високою якістю апаратури, а також тим, що тестові сигнали моноуральні, а бінауральні. [18]

1.1.4. MUSHRA

Інший метод тестування широкосмугових каналів, іменований MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor) (ITU-R Rec. BS.1534-1, 2003 [7]), складається в пред'явленні слухачеві безлічі зразків сигналів, в тому числі зразка еталонного сигналу. Якість сигналу оцінюють за 100-бальною шкалою, від 0 (погано) до 100 (відмінно). Цю шкалу потім ділять на 5 градацій, аналогічно шкалою методу ACR, і остаточне рішення приймають, порівнюючи сигнали з еталонним сигналом [18].

1.2. Об'єктивні методи оцінювання якості сигналів

Об'єктивні методи оцінювання якості мови реалізуються шляхом сегментації мови сигнал у фрейми тривалістю 10-30 мс, а потім обчислюється міру спотворення між оригіналом і обробленим сигналом. Оцінка отримання в результаті розрахунку є усередненим значенням спотворення по всіх фреймах. Як буде видно далі, обчислення вимірювання спотворення може здійснюватися або в часовій області, або в частотній області. Для вимірювань в частотній області передбачається, що будь-які спотворення або розбіжності, виявлені у величині спектра, співвідносяться з якістю сигналу. Огляд об'єктивних заходів можна знайти в [1], [15-21].

Далі приведено опис тих методів, які вже використовувались різними авторами в оцінюванні якості музичних сигналів, про що йшлося у вступній частині дисертації, і будуть використовуватись в експериментальній частині даної роботи.

1.2.1. Segmental Signal to Noise Ratio - SSNR

Сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio) представляє групу показників часової області:

$$SSNR = \frac{1}{M} \sum_{m=1}^M 10 \lg \left[\frac{\sum_{n=R(m-1)+1}^{Rm} x^2(n,m)}{\sum_{n=R(m-1)+1}^{Rm} [x(n,m) - y(n,m)]^2} \right], \quad (1.1)$$

де $x(n,m)$ і $y(n,m)$ - n -ті вибірки m -го фрейма чистого сигналу $x(n)$ і спотвореного сигналу $y(n)$, відповідно; M - кількість фреймів; R - кількість вибірок в фреймі. При обчисленні показника SSNR враховуються тільки ті сегменти, для яких відношення сигнал-шум знаходиться в інтервалі $[-10, +35]$ дБ.

Як показано в [16] SSNR можна використовувати для оцінки якості фільтрованих сигналів. якщо врахувати, що істотним недоліком показника SSNR є його висока чутливість до похибки вирівнювання по часу порівнюваних сигналів.

Ефективним засобом зниження цієї чутливості є підвищення частоти дискретизації, шляхом інтерполяції порівнюваних сигналів, в 2...4 рази. Крім того, порядок фільтра повинен бути непарною для точного вирівнювання сигналів за часом. [16]

1.2.2. Logarithmic Spectral Distortion - LSD

Логарифмічні спектральні спотворення LSD (Logarithmic Spectral Distortion) належить до групи показників частотної області:

$$LSD = \frac{2}{MJ} \sum_{m=1}^M \sum_{j=1}^J |G\{X(j,m)\} - G\{Y(j,m)\}|, \quad (1.2)$$

$$G\{X(j,m)\} = \max\{20 \lg(|X(j,m)|), \delta\},$$

$$\delta = \max_{l,k} \{20 \lg(|X(l,k)|)\} - 50$$

де $X(j,m)$ і $Y(j,m)$ - дискретні перетворення Фур'є m -го фрейма сигналів $x(n)$ і $y(n)$, відповідно, j - номер частотної вибірки, J - кількість частотних вибірок.

На відміну від SSNR, показник LSD значно менш чутливий до похибки вирівнювання сигналів. В роботі [17] це пояснюється тим, що амплітуда спектр сигналу не змінюється при зсуві сигналу за часом. Проте автором зазначається, що показник LSD містить ряд помітних локальних порушень монотонності, які буди виявлені при оцінюванні якості мовних сигналів.

Також автор висуває припущення, щодо причин порушення монотонності. Такими причинами є:

- 1) нормування сигналів по дисперсії;
- 2) чутливість заходи LSD до наявності формант в спектрі мовного сигналу;
- 3) статистична нестійкість оцінки LSD.

Врахування перерахованих факторів дозволяє підвищити достовірність оцінювання заходи LSD. [17]

1.2.3. Bark Spectral Distortion - BSD

Барк-спектральне спотворення BSD (Bark Spectral Distortion) був розроблений Веном (Wang) у 1992 році і відноситься до групи перцептивних показників, які враховують особливості слуховий системи людини. В методі присутня емуляція перцептивних перетворень кількох слухових явищ, таких як інтеграція критичної смуги в равлику (обмежена першими 15 блоками в діапазоні 50-3400 Гц) і стиснення гучності. Алгоритм розраховує відмінність для кожного фрейма тривалістю 10 мс з перекриттям фреймів на 50%. у спектрі еталонного та спотвореного сигналу за шкалою барків. Математичне представлення алгоритму є наступним[21]:

$$BSD = \frac{\sum_{m=1}^M \sum_{k=1}^K [B_x(k, m) - B_y(k, m)]^2}{\sum_{m=1}^M \sum_{k=1}^K [B_x(k, m)]^2}, \quad (1.3)$$

$B_x(k, t)$ і $B_y(k, t)$ - барк-спектри m -го фрейма сигналів $x(n)$ і $y(n)$, відповідно, k - номер критичної смуги частот.

Базовим тут є поняття барк-спектра, обчислення якого проводяться поетапно:

- обчислюють спектр потужності сигналу $P(\omega)$;
- шкалу частот перетворюють в шкалу барків Ω ;
- формують АЧХ $\Psi(\Omega)$ барківського фільтра;
- спектр потужності $P(\Omega)$ згортають з АЧХ $\Psi(\Omega)$ барківського фільтра;
- отриманий результат множать на криву рівної гучності $E(\omega)$;
- коректують шкалу гучності, витягуючи корінь кубічний з результату попереднього етапу (перехід з фонів в сони).

У дослідженні Вена, метод показав відносно високий коефіцієнт кореляції між прогнозованою якістю оцінки, MOSLQON, і оцінки слухових показників якості, MOSLQSN ($\rho = 0,85$) для алгоритму BSD.

1.2.4. Perceptual Evaluation of Audio Quality - PEAQ

Є дві версії алгоритму PEAQ; «базова» версія використовується у програмах, де обчислювальна потужність є не великою, «розширена» версія є більш перцептивно точною, ніж «базова», але в чотири рази більш вимоглива до обчислювальної потужності і використовується там, де точність є надзвичайно важливою. [22]

Основною структурною відмінністю між «базовою» та «розширеною» версіями є те, що «базовою» версія має лише одну периферійну модель вуха на базі ШПФ (рис.??), а в «розширеній» версії дві периферійні моделі вуха: на

Як відомо у зовнішньому та середньому вусі, явно присутні резонансний та фільтруючий ефекти коли звукові хвилі перетворюються на механічні коливання на барабанній перетинці. Три крихітні кістки (молоток, коваделко і стремено) виступають як трансформатор між зовнішнім вухом з повітряним середовищем і рідиною, якою наповнене внутрішнє вухо. Відбувається, по суті, узгодження імпедансу для забезпечення мінімальної втрати енергії через відбиття. Алгоритм PEAQ намагається моделювати характеристики впливу зовнішнього та середнього вуха на звукові сигнали використовуючи підхід Терхардта, який моделює їх ефекти, включаючи внесок внутрішнього шуму в вусі. Частина частотної характеристики представлена на рис.1.3 показує, що зовнішнє середнє вухо діє як смуговий фільтр з резонансом між 2 і 4 кГц (близько 3 кГц). [27]

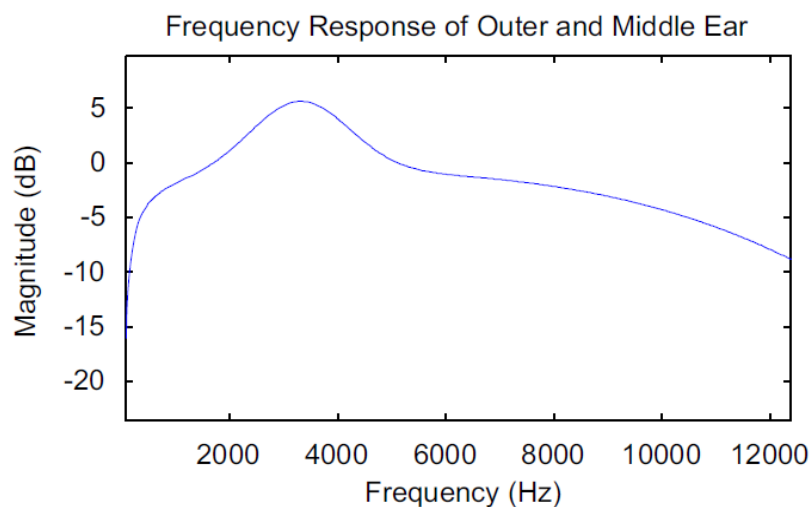


Рис.1.3. Інтегральна АЧХ зовнішнього та внутрішнього вуха [27]

У равлику внутрішнього вуха волоскові клітини є рецепторами звукового тиску. Відбувається перетворення частоти відповідно до положення клітин і положення максимуму збудження залежить від частоти вхідного сигналу. Кожна точка уздовж базилярної мембрани пов'язана з конкретною характеристичною частотою (критичною частотою). Ширина критичної смуги визначена Цвикером, коливається від верхньої частоти зрізу 100-15500 Гц ,

тобто 24 барка = 15500 Гц. Шкала частот, що використовується в REAQ, - це варіація попередньої і коливається від 80Гц-18кГц діапазони відрізняється для моделі на базі ШПФ, що використовуються в «баовій» та «розширеній» версіях. Роздільна здатність 0.25 барка використовується в «базовій» версії в той час як у «розширеній» версії використовується роздільна здатність 0.5 барка. [22]

Ці смуги пропускання барківської частоти ведуть до 109 критичних смуг фільтрів в моделі вуха на основі ШПФ в «базовій» версії та 55 критичних смуг для моделі на базі ШПФ, що використовується в «розширеній» версії. У REAQ частотні компоненти, вироблені ШПФ (зважені за частотною характеристикою зовнішнього та середнього вуха) згруповані у смуги частот, як це в слуховій системі людини. Енергія ШПФ у кожній критична смузі підсумовується для створення єдиного енергетичного значення для кожної смуги. Наступний крок для моделі вуха на основі ШПФ - це додавання частотно залежного зміщення для кожної критичної смуги, як показано на рис.1.4. Зсув репрезентує внутрішній шум, вироблений всередині людського вуха. Внутрішній шум - це виразний маскер, який виробляє постійний поріг маскування, більш широко відомий як "поріг слуху у тиші". Стандарт REAQ описує сигнали в цьому пункті, як Шаблони Висоти Тону. [27]

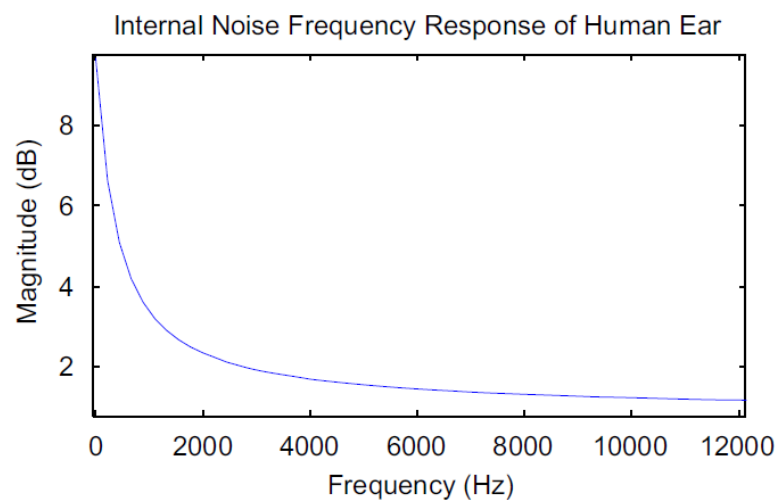


Рис.1.4. Спектр внутрішнього шуму вуха [27]

Шаблони висоти тону розмиваються за частотою використовуючи функцію розповсюдження, що залежить від рівня, яка моделює одночасне маскування. В нижній перехідній зоні спад - постійний 27 дБ / Барк, як показано в (1.5). Тіед (Thiede)[23], який розробив алгоритм оцінки якості звуку DIX, на якому ґрунтується багато частин PEAQ, вказує на те, що під час експериментів зміна швидкості нижнього спаду не мала істотного впливу на ефективність оцінки якості. Тіед використовував найвище значення швидкості спаду знайденого в літературі, яке становить 31 дБ / Барк. Проте верхній спад, який використовується в PEAQ, залежним від рівня і частоти (1.4) і (1.5). [27]

$$S_u[k, L(k, n)] \left(\frac{dB}{Bark} \right) = -24 - \left(\frac{230Hz}{f_{C_k}} \right) + 0.2 \cdot \left(\frac{L(k, n)}{dB} \right), \quad (1.4)$$

$$S_l[k, L(k, n)] = 27 \left(\frac{dB}{Bark} \right), \quad (1.5)$$

де L - шаблон Висоти тону, f_c - центральна частота, k - індекс критичного діапазону і n – номер фрейма. S_u - розрахунок верхніх схилів, а S_l - розрахунок нижніх схилів. [27]

Розповсюдження (маскування) здійснюється незалежно для кожної критичної смуги та результати називаються в стандарті Незмінні Шаблони Збудження.

В «базовій» версії PEAQ з ефектів тимчасового маскування використовується лише маскування вперед, вплив на модель вуха на базі ШПФ зворотного маскування характеристик тимчасової маскування ефекти як роздільна здатність з точки зору загальної продуктивності незначний. Зазвичай маскування триває лише кілька (зазвичай 5-10) мілісекунд, в той час як фрейми в PEAQ мають довжину приблизно 21 мс. Маскування вперед моделюється як простий фільтр нижньої частоти першого порядку, який використовується щоб маскувати енергії в кожній критичній смузі з часом. [28]



Рис.1.5. Структурна схема «розширеної» версії показника PEAQ

У «розширеній» версії PEAQ модель вуха на основі банку фільтрів використовується разом із моделлю вуха на основі ШПФ (рис.1.5). У моделі вуха на основі банку фільтрів обробка здійснюється в часовій області, а не у коротких фреймах, як у моделі вуха на базі ШПФ. Модель вуха на базі банку фільтрів забезпечує більш точне моделювання людського вуха, оскільки воно використовує більш точну часову роздільну здатність, отже, стає можливим моделювання зворотного маскування і тимчасової тонкої структури сигналу (відчуття шорсткості). [28]

Рівень прослуховування 92 дБ встановлюється, як і в моделі вуха на основі ШПФ. Еталонний і деградований сигнали обробляються індивідуально. Різні під-вибірки реалізовані для зменшення обчислювальних зусиль на різних етапах обробки. Сигнали розбиваються на смугові сигнали банком фільтрів, що містить критичні смуги, розташовані на однаковій відстані. Банк фільтрів має 40 фільтрів з центральними частотами від 50 Гц до 18 кГц, центральна частоти розташовані рівномірно на шкалі барків. Кожна критична смуга складається з двох фільтрів з однаковою частотною характеристикою, але один з них має 90° фазовий зсув (перетворення Гільберта). Обвідні їх імпульсних

відгуків мають форму вікна Ханнінга (\sin^2). Коефіцієнти FIR фільтрів можуть бути розраховані, використовуючи наступні рівняння: [27]

$$\begin{aligned} h_{re}(k, n) &= \frac{4}{N[k]} \sin^2 \left(\pi \frac{n}{N[k]} \right) \cos \left(2\pi f_c[k] \left(n - \frac{N[k]}{2} \right) T \right) \\ h_{im}(k, n) &= \frac{4}{N[k]} \sin^2 \left(\pi \frac{n}{N[k]} \right) \sin \left(2\pi f_c[k] \left(n - \frac{N[k]}{2} \right) T \right) \end{aligned} \quad \begin{matrix} 0 \leq n \leq N[k] \\ \\ \end{matrix}$$

(1.6)

$$h_{re}(k, n) = h_{im}(k, n) = 0 \quad | \quad n < 0, n \geq N[k]$$

де k – індекс критичної смуги від 1 до 40, n – це номер вибірки і T - час вибірки в секундах.

Після банку фільтрів, наступна частина алгоритму моделює фільтраційний ефект зовнішнього та середнього вуха що виконується так само, як в моделі на базі ШПФ. Одночасне маскування (поширення частоти) також моделюється, як у моделі вух на основі ШПФ. Потім миттєва енергія кожного виходу банку фільтрів розраховується попередньо до тимчасового маскування. Хоча тимчасова маскування вперед реалізується як в моделі на основі ШПФ, так і в моделі на базі банку фільтрів, зворотнє маскування здійснюється тільки моделі на снові в банку фільтрів. FIR фільтр 12-го порядку використовується для моделювання зворотнього маскування. Фільтр розмазує енергію частотного розповсюдження за часом відповідно до (1.7): [27]

$$E_1[k, n] = \frac{0.9761}{6} \sum_{i=0}^{11} E_0[k, n-i] \cos^2 \left(\pi \frac{(i-5)}{12} \right) \quad (1.7)$$

де k - індекс критичної смуги, n - індекс кадру, i – це номер вибірки затримки та E_0 - вихідний сигнал фільтра банку енергії. 0.9761 - це константа, яка приймає відтворення враховуючи рівень, а коефіцієнт 6 - швидкість зниження частоти.

Модель вуха, заснована на базі банку фільтрів, завершується включенням внутрішнього шуму та моделювання маскування вперед. Знову ж таки, вони засновані на тому ж принципі, такі як ті, що використовуються в

моделі вуха на основі ШПФ. Структура вихідного шаблону фільтра після маскування та додавання внутрішнього шуму називають "шаблонами збудження". [26]

Когнітивна модель в PEAQ моделює когнітивну обробку людського мозку, яка використовується для надання аудіосигналу якісної оцінки. У PEAQ когнітивна модель обробляє параметри, вироблені психоакустичними моделями вуха для формування вихідних параметрів, відомі як MOV та проводить картування MOV до єдиної оцінки ODG.

Величина шкали ODG залежить від значення опорних точок п'ятибальної шкали знецінення. Значення цих опорних точок пов'язано з суб'єктивним визначенням якості, і може змінюватися з часом. З цієї причини технічний показник якості повинен бути виражений не як оцінка різниці, а як більш абстрактна одиниця, яка монотонно відображає ODG. Якщо опорні точки ODG зміняться, ця міра залишиться незмінною, і тільки відображення ODG має бути скориговане.

Зручним способом отримати такий значення є використання остаточної нелінійності вихідного шару нейронна мережа. На даному етапі всі MOV вже об'єднані в єдине значення, але остаточно масштабування до діапазону SDG ще не відбулось. Це значення називається індекс спотворення (Distortion Index - DI). Вхідні дані (MOV) до нейронної мережі пов'язані з DI, наступним виразом: [27]

$$DI = w_y[J] + \sum_{j=0}^{J-1} \left(w_y[j] \operatorname{sig} \left(w_x[l, j] + \sum_{i=0}^{I-1} w_x[I, j] \frac{x[i] - a_{\min}[i]}{a_{\max}[i] - a_{\min}[i]} \right) \right) \quad (1.8)$$

У вищезгаданому рівнянні термін x представляє вхідні дані (MOV). Термін sig означає сигмоїдну функцію активації Вагові коефіцієнти для вхідних та вихідних значень називаються w_x і w_y , відповідно, і даються в BS.1387 [22]. Вони були розраховані використанням даних суб'єктивного оцінювання. Рівняння для обчислення ODG з DI є наступним: [27]

$$ODG = b_{\min} + (b_{\max} - b_{\min}) \operatorname{sig}(DI) \quad (1.9)$$

де b_{min} , b_{max} є попередньо визначеними масштабними коефіцієнтами, DI - індекс спотворення.

Вихідні масштабні коефіцієнти b_{min} та b_{max} наведені в стандарті [22], який, однак, не деталізує як вони були отримані. Термін "sig" означає "сигмоїдну" функцію активації. Шкала оцінки якості аудіо-сигналу ODG коливається від 0 до -4, де 0 - оптимальна якість і -4 - дратуючі спотворення.

Алгоритм був широко випробуваний в ході його розвитку, з широким спектром звукових сигналів різних типів, включаючи джаз, рок, тубу, мову тощо, інструментами такими як трикутники, кларнети, клаве, арфи, барабани, саксофони, волинками тощо. Звукові сигнали були високоякісними і спотворювались впливом таких кодеків: MPEG1, NICAM, Dolby та Mini Disc. Деякі аудіо матеріали, що були використані, оброблись каскадом кодеків, і деякі з них містять помилки квантування, нелінійні спотворення і шум. Кожен сигнал був тривалістю від 10 до 20 секунд. Оцінки якості за алгоритмом (об'єктивні) порівнювались з результатами, отриманими з тестів аудіювання, з яких було встановлено, що коефіцієнт кореляції між об'єктивним та суб'єктивним оцінками складала 0,837 для «базової» версії і 0,851 для «розширеної» версії. [27]

1.3. Висновки

У даному розділі приведено суб'єктивні та об'єктивні методи оцінки якості сигналів, що можуть використовуватись, як для оцінки мови так і музики, описано їх суттєві ознаки та особливості. На основі цих даних обрано методи для проведення експериментального дослідження якості музичних сигналів.

Для суб'єктивного тестування обрано метод, іменований «оцінкою ступеня деградації» (Degradation Category Rating - DCR). Відповідно до цього методу, аудитори на слух порівнюють якість спотвореного і еталонного

мовних сигналів, виставляючи оцінки за п'ятибальною шкалою, іменованою «Шкалою середньої деградації» (Degradation Mean Opinion Score - DMOS).

Цей показник був обраний з огляду на те, що на відміну від метода ACR, він використовує при прослуховуванні порівняння з еталонним сигналом. Такий підхід дозволяє більш точно аудиторам дати кількісну оцінку ступені деградації спотвореного сигналу. Від показників MUSHRA та Оцінки малих спотворень, DCR якісно відрізняється простотою. Немає сенсу використання складного методу Оцінки малих спотворень, оскільки зменшення смуги пропускання до 10 кГц і менше повинна суттєво відображатись на якості сигналу.

В якості об'єктивних заходів якості музичного сигналу обрано всі чотири показники, такі як: сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio), логарифмічне спектральне спотворення LSD (Logarithmic Spectral Distortion) – через їх простоту в обчисленні; барк-спектральне спотворення BSD (Bark Spectral Distortion) і перцептивна оцінка якості звуку PEAQ (Perceptual Evaluation Audio Quality) – оскільки вони враховують особливості слухової системи людини і мають високі показники кореляції з суб'єктивними оцінками. Однак слід зауважити, що метод PEAQ обраний для дослідження в «базовій» версії алгоритму, через начну вимогливість до ресурсів комп'ютера «розширеної» версії.

Варто пам'ятати що методи оцінювання SSNR, LSD та BSD є монофонічними, що необхідно додатково враховувати при використанні їх щодо музичних сигналів. PEAQ на відміну від них, може працювати і зі стререофонічними записами, та має іншу особливість – частота дискретизації файлів що оцінюються має складати 48 кГц.

РОЗДІЛ 2. ОРГАНІЗАЦІЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

Як вже було сказано раніше основною метою даної дисертації є виявлення смуги пропускання телекомунікаційного тракту достатнього для якісної передачі музичних сигналів. Також виявлення залежностей між об'єктивними та суб'єктивними показниками та аналіз отриманих даних.

Будь-яке дослідження для отримання достовірної інформації вимагає, щоб дослідження проводились в контрольованих умовах. Для досягнення цього оцінювання якості музичних сигналів було проведено за наступною схемою (рис.2.1), де $x(t)$ - еталонний (неспотворений) фрагмент музичного сигналу, а $y(t)$ – спотворена, за допомогою ФНЧ, копія фрагмента музичного сигналу.



Рис.2.1. Схема експериментального дослідження

Послідовність випробувань включає в себе два етапи: підготовка (підбір музичного матеріалу, розробка програмного забезпечення для автоматизації процедури) і власне проведення експерименту (прослуховування, отримання оцінки, обробка експериментальних даних).

2.1. Підготовчий етап

В якості еталонних музичних сигналів було обрано 8 музичних творів, по 2 твори для кожного чотирьох жанрів: класична музика, поп-музика, джаз та рок-музика відповідно.

Твори обирались так , щоб навіть в рамках одного жанру вони значно відрізнялись за характером, а також мали значне різноманіття та присутність музичних інструментів в рамках жанру. Якість обраних творів відповідала стандарту Audio CD з частотою дискретизації 44100 Гц і бітовою глибиною 16 біт.

Нижче приведені назви обраних музичних творів:

1. Зразки класичної музики
 - **classic1:** П.І. Чайковський – *Пори року. Серпень*
 - **classic2:** Johannes Brahms – *Hungarian Dance #5*
2. Зразки Поп-музики
 - **pop1:** LP – *Lost on you*
 - **pop2:** Karmin – *Dance with you*
3. Зразки Рок-музики
 - **rock1:** System of a down – *Question!*
 - **rock2:** Kings of Leon – *Sex on fire*
4. Зразки музики в стилі джаз
 - **jazz1:** Avishai Cohen – *Nu Nu*
 - **jazz2:** Laco Tayfa – *Atamaca*

Для оцінювання якості музичних сигналів було обрано один суб'єктивний та чотири об'єктивних показника, кожен із них має свої особливості детально описані в розділі 1, які треба враховувати при проведенні оцінювання. Загалом же суть всіх обраних методів зводиться до порівняння еталонного сигналу зі спотвореним для отримання кількісної оцінки якості сигналу.

Спотворення музичного сигналу проводилося шляхом почергового пропускання його через нерекурсивні фільтри низької частоти, синтезовані за методом Ремеза. Характеристики фільтрів:

- частота зрізу від 1 до 20 кГц з кроком 1 кГц;
- розмір перехідної зони 5% від смуги пропускання;

- нерівномірність АЧХ в смузі пропускання 1 дБ;
- загасання -70 дБ

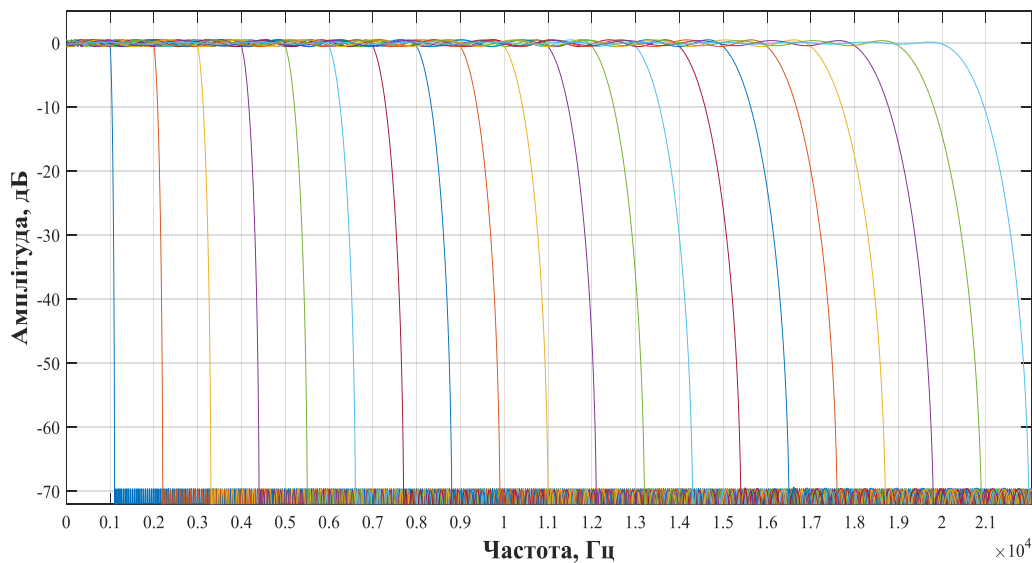


Рис.2.2. АЧХ нерекурсивних НЧ фільтрів синтезованих за методом Ремеза

Для полегшення проведення та обробки даних оцінювання в програмному середовищі MatLab було створено дві програми.

Перша програма відтворювала по чергово обраний аудитором еталонний і спотворений, за допомогою ФНЧ описаного вище, фрагменти музичних сигналів. Тривалість фрагменту обирається слухачем з проміжку 15...30 секунд. Смуга пропускання фільтру генерується програмою у випадковій послідовності, щоб виключити можливість коригування оцінок аудитором відносно згенерованої смуги пропускання. Аудитор прослуховуючи ці сигнали, відповідно до методу іменованого «оцінкою ступеня деградації» (Degradation Category Rating - DCR), порівнює їх і вносить в програму оцінки за п'ятибальною шкалою, іменованою «Шкалою середньої деградації» (Degradation Mean Opinion Score - DMOS). В кінці тестування програма генерує персональний файл аудитора, який включає в себе оцінки якості сигналів відповідно до смуги пропускання фільтра.

Друга програма займається підготовкою музичних фрагментів довжиною 40 секунд для подальшої обробки їх підпрограмами, що виконують

оцінювання якості відповідно до обраних об'єктивних показників. В якості об'єктивних методів оцінки якості музичного сигналу використані чотири показники, таких як сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio), логарифмічне спектральне спотворення LSD (Logarithmic Spectral Distortion), барк-спектральне спотворення BSD (Bark Spectral Distortion) і перцептивна оцінка якості звуку PEAQ (Perceptual Evaluation Audio Quality). Кожен із них має свої особливості описані в розділі 1, що й призвело до потреби створення програми.

Оскільки показники SSNR, LSD, BSD в першу чергу для оцінки якості мови спотворених каналами зв'язку, тому вони працюють лише з моно-сигналами. З огляду на це програма перетворює музичні фрагменти еталонного сигналу із стерео в моно. Після цього по черзі фільтрує отримані музичні фрагменти за допомогою вище згаданого ФНЧ, вирівнює еталонний сигнал і спотворену копію в часі, і тоді відправляє їх на обробку підпрограмам оцінювання якості.

Алгоритм PEAQ розроблявся спеціально для оцінювання якості музики, тому перетворення зі стерео в моно йому не потрібне. Однак він має іншу особливість – працює лише з файлами частота дискретизації яких 48 кГц. Виходячи з цього після фільтрації еталонний та спотворений сигнали піддаються передискретизації з 44,1 кГц в 48 кГц і тільки після цього відправляються на обробку алгоритмом PEAQ.

2.2. Проведення експерименту

Суб'єктивне оцінювання якості музичного сигналу проводилося 23 аудиторами, юнаками і дівчатами у віці 19-24 років без явних вад слуху, із використанням фрагментів сигналів тривалістю від 15 до 30 секунд. Для прослуховувань використовувався комп'ютер зі встановленим пакетом MatLab і навушники.

Для прослуховування обов'язковим було використання навушників з хорошими ізоляційними характеристиками для нівелювання впливу на результати оцінювання можливого навколишнього шуму та реверберації приміщення. Частотний діапазон навушників має відповідати смузі 20 - 20000 Гц або більше.

Об'єктивне оцінювання проводилось автоматизовано програмою розробленою в середовищі MatLab з використанням підпрограм, що відповідають алгоритмам розрахунку обраних показників якості.

2.3. Висновки

В даному розділі описано ключові моменти організації експериментального дослідження для виявлення залежності якості музики від смуги пропускання. При підготовці до експерименту враховувались особливості кожного методу оцінювання для отримання найбільш достовірних даних.

Для зручності та полегшення обробки отриманої інформації було створено допоміжні програми, які враховують особливості кожного методу оцінювання.

Описано вимоги щодо підбору музичного матеріалу для прослуховування, аудиторів та необхідної апаратури.

РОЗДІЛ 3. АНАЛІЗ РЕЗУЛЬТАТІВ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

3.1. Результати суб'єктивного та об'єктивного оцінювання

На рис.3.1 представлені усереднені результати суб'єктивного оцінювання за шкалою DMOS. Спостерігається монотонне зростання якості сигналу із розширенням смуги частот до 14кГц, подальше розширення смуги частот не призводить до зростання якості музичного сигналу. Також на графіку червоним пунктиром показана дисперсія оцінки з 95% довірчим інтервалом. Такий результат є не зовсім очікуваним, оскільки у вже згаданій роботі [17], було досить багато прогалин, що не давало підстав значно покладатися на отриманий авторами результат в 14-15 кГц. Однак їх результати підтвердились і добре корелюють з гіпотезою висловленою вже згадуваним Бранденбургом у [16].

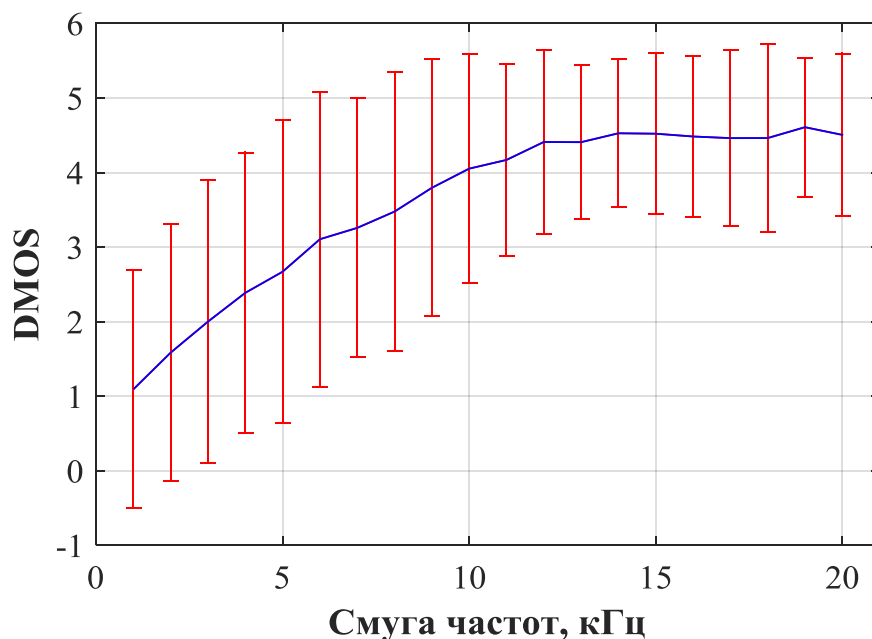


Рис.3.1. Результати суб'єктивного оцінювання за шкалою DMOS

Усереднені результати об'єктивного оцінювання якості цих сигналів можна спостерігати на рис.3.2. В залежності від методу значення верхньої

границі смуги частот, після якої підвищення якості не спостерігається, варіюється в мажах 12-14 кГц.

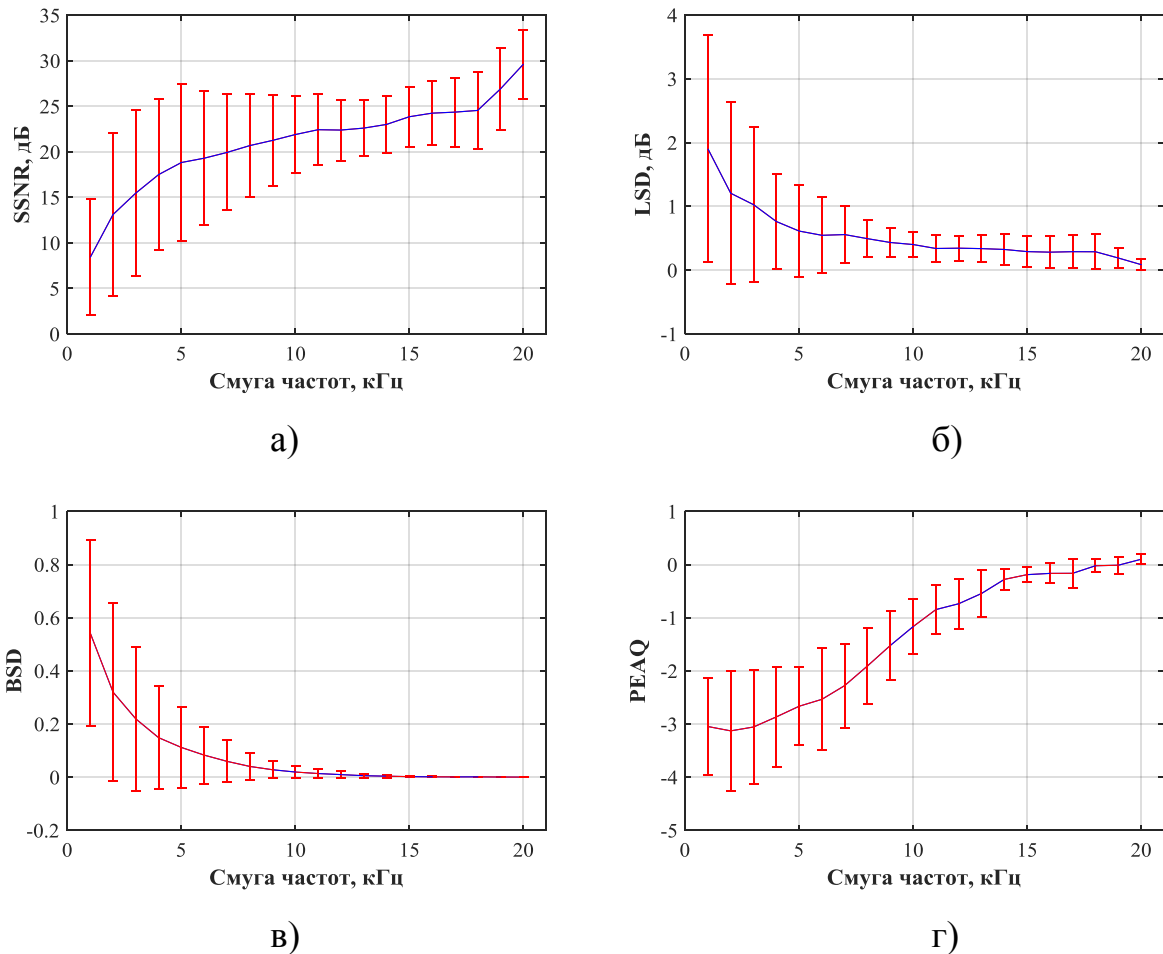


Рис.3.2. Усереднені результати об'єктивного оцінювання за показниками
а) SSNR, б) LSD, в) BSD, г) PEAQ.

На графіку оцінки SSNR можна помітити різке підвищення якості сигналу при розширенні смуги понад 18 кГц, схожий артефакт можна побачити у даного метода для всіх музичних файлів на рис.3.3. Такий ефект пояснюється залежністю даного показника від частоти дискретизації, що зазначено в [16]. Усунути цей недолік можливо за допомогою підвищення частоти дискретизації в 2-4 рази, однак в даному дослідженні такий підхід не був використаний, оскільки для цього не вистачало оперативної пам'яті комп'ютера.

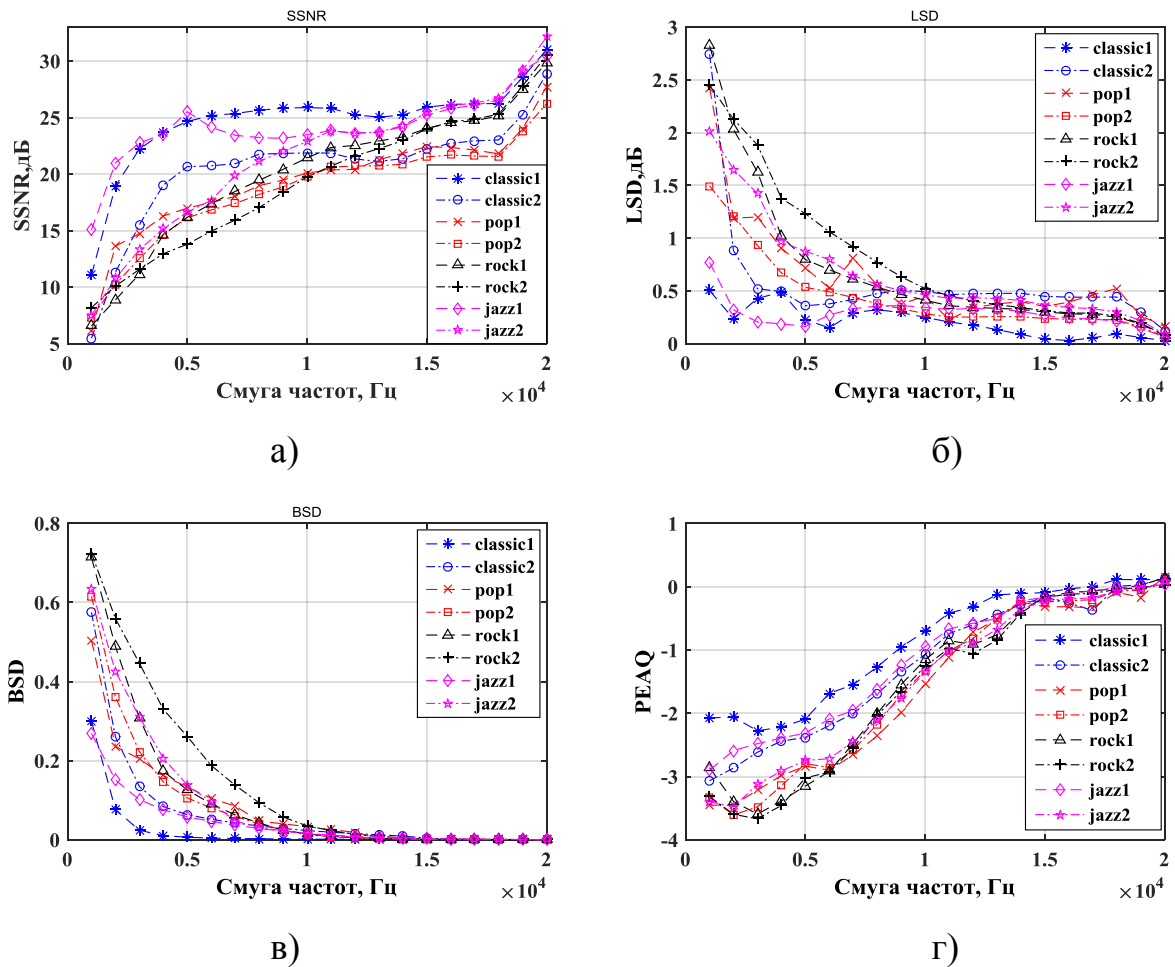


Рис.3.3. Результати об'єктивного оцінювання за показниками а) SSNR, б) LSD, в) BSD, г) PEAQ.

Також слід звернути увагу на графік показника LSD, в [13], [15] та [17] показано високу чутливість до особливостей спектру досліджуваних музичних сигналів, що призводить до порушення монотонності цього показника, як і для деяких музичних фрагментів представлених на рис.3.3. В [13] порушення монотонності графіків показника LSD пов'язують із тим, що АЧХ фільтрів не є ідеальною. Однак на рис.3.2. значних порушень монотонності показаних у вищезгаданих роботах не спостерігаються, що можна пояснити значною мірою вибором музичних зразків для тестування. При достатній різноманітності музичного матеріалу вищезгадані недоліки нівелюються.

На графіку оцінки PEAQ є порушення монотонності для деяких музичних файлів у смугах частот менше 3 кГц, яка відображається і на усередненому графіку. Однак зважаючи на те, що алгоритм розроблявся для порівняння сигналів з малими спотвореннями, на чому наголошують в [27], такий результат є логічним оскільки смуга частот в 3 кГц є надто вузькою для музики і відмінність сигналу від еталонного буде значною.

3.2. Співставлення результатів суб'єктивного та об'єктивного оцінювання

Для оцінки надійності мір якості сигналів зазвичай розраховують значення коефіцієнтів кореляції, що зроблено і в даній роботі. Коефіцієнти кореляції між результатами об'єктивного і суб'єктивного оцінювання, наведені в табл.3.1.

Таблиця 3.1. Коефіцієнти кореляції об'єктивних показників

Об'єктивний показник	SSNR	LSD	BSD	PEAQ
Коефіцієнт кореляції	0.94	-0.93	-0.92	0.94

Як можна побачити всі чотири методи показали високе значення коефіцієнта кореляції. Схожий підхід для оцінки надійності мір якості використовується майже у всіх виданнях на цю тематику, і зазвичай на цьому етапі всі дослідники зупиняються.

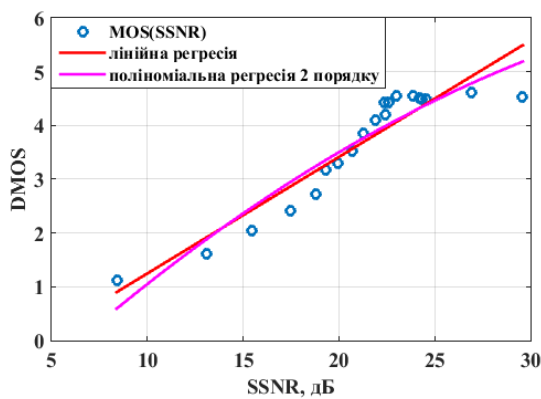
Однак в роботі [15] приводяться аргументи, що такий підхід не дуже коректний, оскільки між результатами об'єктивного і суб'єктивного оцінювання може існувати нелінійна залежність, а коефіцієнт кореляції є мірою лінійної залежності випадкових величин. Виходом із ситуації може служити побудова «карти» усередненої залежності між результатами об'єктивних і суб'єктивних вимірів якості сигналів.

Такий підхід використовують досить давно, його іноді називають «картуванням» [15] або, інакше кажучи, приведенням результатів об'єктивних

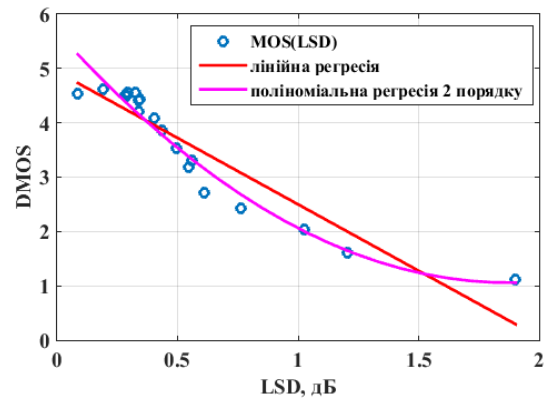
вимірювань до шкали результатів суб'єктивних вимірів. Сутність картування полягає в первісному побудові залежностей значень суб'єктивного і об'єктивного показника від деякого важливого для даної ситуації параметра, що визначає ступінь спотворень.

На рис.3.4 представлено лінії регресії відповідності об'єктивний оцінок суб'єктивній, так звані «карти відповідності». Побудова наведених карт відповідності, виконана двома способами: з допомогою лінійної і поліноміальної регресії.

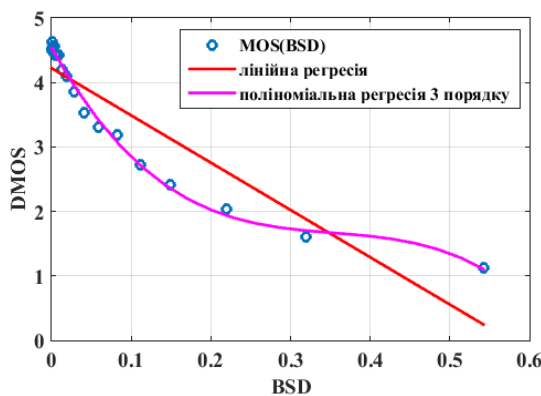
Оптимальним слід вважати вибір апроксимуючого полінома мінімального порядку, при якому досягається прийнятна для методів помилка апроксимації і не порушується монотонний характер апроксимації.



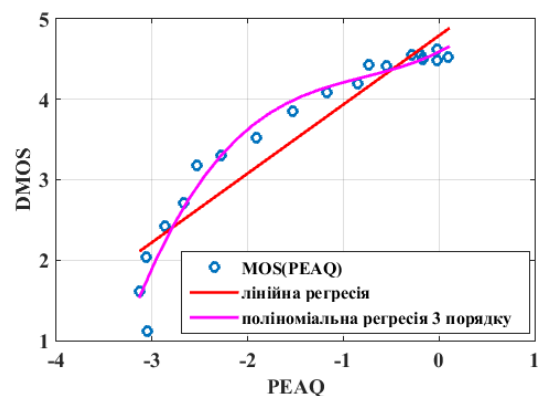
а)



б)



в)



г)

Рис.3.4. Апроксимація залежності між результатами об'єктивного і суб'єктивного оцінювання якості музичних сигналів: а) DMOS-SSNR; б) DMOS-LSD; в) DMOS-BSD; г) DMOS-PEAQ

Розраховано також максимальні похибки апроксимації для перших п'яти порядків поліному. Значення розрахованих похибок представлені в таб.3.2.

Таблиця 3.2. Максимальна похибка апроксимації об'єктивних показників

Порядок полінома	Об'єктивний показник			
	SSNR(дБ)	LSD(дБ)	BSD	PEAQ
1	1.071	4.239	4.284	0.994
2	1.069	4.140	3.500	0.585
3	1.616	3.907	3.425	0.427
4	1.552	3.470	3.652	0.471
5	1.494	4.500	3.461	0.491

Як можна помітити, не зважаючи на високий коефіцієнт кореляції у всіх чотирьох показників якості, для всіх є ярко вираженою нелінійний характер залежності.

Найбільш неочікуваним є той факт, що для алгоритму PEAQ, який на даний момент є найбільш точним інструментом в оцінюванні якості музичних сигналів, і враховує значну кількість відомих перцептивних явищ, притаманний, як і для інших розглянутих мір якості, нелінійний характер залежності. Всі представлені раніше роботи [11], [26], [27] та [29], з використанням даного алгоритму для оцінювання якості від різноманітних параметрів (глибини квантування, щільності бітового потоку) зазвичай мали невелику представницькість повноцінних музичних творів, зазвичай обирались окремі музичні інструменти, що як бачимо не є досить показовим, через те, що більшість музичних інструментів мають не дуже широкий частотний діапазон. Також всі вони завершували свої дослідження на етапі кореляційного аналізу. Такий результат, ще раз доказує, що до результатів кореляційного аналізу слід відноситись з обережністю. А підбір тестового

матеріалу є важливою складовою підготовчого процесу до проведення експерименту.

3.3. Висновок

Шляхом експериментальних досліджень отримані суб'єктивні та об'єктивні оцінки якості музичних сигналів в залежності від смуги пропускання для достатньо широкого набору тестових сигналів. Виявлено граничне значення смуги в пропускання в 15 кГц.

Для показника LSD показано, що при достатній представницькості музичного матеріалу підбраного для тестування можна нівелювати вплив чутливості даного показника до спектральних компонент сигналу.

Проведено зіставлення результатів суб'єктивних та об'єктивних показників Показано нелінійний характер залежності між показником PEAQ в базовій версії та DRC.

В результаті оцінки коефіцієнта кореляції та побудови карт відповідності між суб'єктивними та об'єктивними оцінками показано перевагу у використанні карт відповідності замість коефіцієнта кореляції.

РОЗДІЛ 4. РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ

4.1 Опис ідеї проекту

Таблиця 1. Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
	1.	
	2.	

Таблиця 2. Визначення сильних, слабких та нейтральних характеристик ідеї проекту

№ п/ п	Техніко економічні характерист ики ідеї	Потенційні товари/концепції конкурентів				Слабк а сторо на (W)	Нейтраль на сторона (N)	Сильн і сторо на (S)
		Мій прое кт	Конкуре нт 1	Конкуре нт 2	Конкуре нт 3			

4.2 Технологічний аудит ідеї проекту

Таблиця 3. Технологічна здійсненність ідеї проекту

№ п/п	Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій

4.3 Аналіз ринкових можливостей запуску стартап-проекту

Таблиця 4. Попередня характеристика потенційного ринку стартап-проекту

№ п/п	Показники стану ринку(найменування)	Характеристика
1	Кількість головних гравців, од	
2	Загальний обсяг продаж, грн./ум.од	
3	Динаміка ринку(якісна оцінка)	Зростає/спадає/стагнує
4	Наявність обмежень для входу (вказати характер обмежень)	
5	Специфічні вимоги до стандартизації та сертифікації	
6	Середня норма рентабельності в галузі (або по ринку),%	

Таблиця 5. Характеристика потенційних клієнтів стартап-проекту

№ п/п	Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
	Базова потреба, яку задовольняє товар (згідно концепції потенційного товару)	Визначити потенційні цільові групи клієнтів, що можуть бути зацікавлені у задоволенні означеної потреби	Вписати фактори, що формують поведінку клієнта (стандарти, технічні регламенти, інші фактори цінового та нецінового характеру) та особливості купівлі та експлуатації товару	- до продукції - до компанії-постачальника

Таблиця 6. Фактори загроз

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії

Таблиця 7. Фактори можливостей

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії

Таблиця 8. Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
1. Вказати тип конкуренції-олігополія		
2. За рівнем конкурентної боротьби - світовий		
3. За галузевою ознакою - внутрішньогалузева		
4. Конкуренція за видами товарів: - товарно-предметна		
5. За характером конкурентних переваг - нецінова		
6. За інтенсивністю		

- марочна/не марочна		
----------------------	--	--

Таблиця 9. Аналіз конкуренції в галузі за М. Портером

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	Навести перелік прямих конкурентів	Визначити бар'єри входу в ринок	Визначити фактори сили постачальників	Визначити фактори сили споживачів	Фактори загроз з боку замінників
Висновки:	Визначити інтенсивність конкурентної боротьби з боку прямих конкурентів	- чи є можливо-сті входу в ринок? - чи є потенційні конкуренти? Строки виходу їх на ринок?	Чи постачальники диктують умови роботи на ринку? Які?	Чи клієнти диктують умови роботи на ринку? Які?	Обмеження для роботи на ринку через товари замінники

Таблиця 10. Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)

Таблиця 11. Порівняльний аналіз сильних та слабких сторін «назва проекту»

№ п/п	Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні з ... (назва підприємства)						
			-3	-2	-1	0	+1	+2	+3

Таблиця 12. SWOT- аналіз стартап-проекту

Сильні сторони:	Слабкі сторони:
Можливості:	Загрози:

Таблиця 13. Альтернативи ринкового впровадження стартап-проекту

№ п/п	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації

4.4. Розроблення ринкової стратегії проекту

Таблиця 14. Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи потенційних клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота вхо-ду у сегмент
Які цільові групи обрано:					

Таблиця 15. Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку

Таблиця 16. Визначення базової стратегії конкурентної поведінки

№ п/п	Чи є проект «перешкодою» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки*

Таблиця 17. Визначення стратегії позиціонування

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)

4.5. Розроблення маркетингової програми стартап-проекту

Таблиця 18. Визначення ключових переваг концепції потенційного товару

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)

Таблиця 19. Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові		
I. Товар за задумом	Опис базової потреби споживача, яку задовольняє товар (згідно концепції), її основної функціональної вигоди		
II. Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1.		
	2.		
	Якість: стандарти, нормативи, параметри тестування тощо		
	Пакування		
	Марка: назва організації-розробника + назва товару		
III. Товар із підкріпленням	До продажу		
	Після продажу		
За рахунок чого потенційний товар буде захищено від копіювання:			

Таблиця 20. Визначення меж встановлення ціни

№ п/п	Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу

Таблиця 21. Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту

Таблиця 22. Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення

4.6 Висновки

ВИСНОВОК

У першому розділі приведено суб'єктивні та об'єктивні методи оцінки якості сигналів, що можуть використовуватись, як для оцінки мови так і музики, описано їх суттєві ознаки та особливості. На основі цих даних обрано методи для проведення експериментального дослідження якості музичних сигналів.

Для суб'єктивного тестування обрано метод, іменований «оцінкою ступеня деградації» (Degradation Category Rating - DCR). Відповідно до цього методу, аудитори на слух порівнюють якість спотвореного і еталонного мовних сигналів, виставляючи оцінки за п'ятибальною шкалою, іменованою «Шкалою середньої деградації» (Degradation Mean Opinion Score - DMOS).

Цей показник був обраний з огляду на те, що на відміну від метода ACR, він використовує при прослуховуванні порівняння з еталонним сигналом. Такий підхід дозволяє більш точно аудиторам дати кількісну оцінку ступені деградації спотвореного сигналу. Від показників MUSHRA та Оцінки малих спотворень, DCR якісно відрізняється простотою. Немає сенсу використання складного методу Оцінки малих спотворень, оскільки зменшення смуги пропускання до 10 кГц і менше повинна суттєво відображатись на якості сигналу.

В якості об'єктивних заходів якості музичного сигналу обрано всі чотири показники, такі як: сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio), логарифмічне спектральне спотворення LSD (Logarithmic Spectral Distortion) – через їх простоту в обчисленні; барк-спектральне спотворення BSD (Bark Spectral Distortion) і перцептивна оцінка якості звуку PEAQ (Perceptual Evaluation Audio Quality) – оскільки вони враховують особливості слухової системи людини і мають високі показники кореляції з суб'єктивними оцінками. Однак слід зауважити, що метод PEAQ обраний для дослідження в «базовій» версії алгоритму, через начну вимогливість до ресурсів комп'ютера «розширеної» версії.

В другому розділі описано ключові моменти організації експериментального дослідження для виявлення залежності якості музики від смуги пропускання. При підготовці до експерименту враховувались особливості кожного методу оцінювання для отримання найбільш достовірних даних.

Для зручності та полегшення обробки отриманої інформації було створено допоміжні програми, які враховують особливості кожного методу оцінювання.

Описано вимоги щодо підбору музичного матеріалу для прослуховування, аудиторів та необхідної апаратури.

Шляхом експериментальних досліджень отримані суб'єктивні та об'єктивні оцінки якості музичних сигналів в залежності від смуги пропускання для достатньо широкого набору тестових сигналів. Виявлено граничне значення смуги в пропускання в 15 кГц.

Для показника LSD показано, що при достатній представницькості музичного матеріалу підбраного для тестування можна нівелювати вплив чутливості даного показника до спектральних компонент сигналу.

Проведено зіставлення результатів суб'єктивних та об'єктивних показників. Показано нелінійний характер залежності між показником PEAQ в базовій версії та DRC.

В результаті оцінки коефіцієнта кореляції та побудови карт відповідності між суб'єктивними та об'єктивними оцінками показано перевагу у використанні карт відповідності замість коефіцієнта кореляції.

СПИСОК ПОСИЛАНЬ

1. Cote N. Integral and diagnostic intrusive prediction of speech. - Springer-Verlag Berlin Heidelberg, 2011, ISBN: 978-3-642-18462-8
2. Philipos C. Luisou. Speech enhancement. Theory and practice. Second edition – CRC Press , 2013, ISBN: 978-1-4665-0422-6
3. ITU-T Rec. P.862 - Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs (2001)
4. JM. Valin, K. Vos, T. Terriberry. Definition of the Opus Audio Codec, IETF RFC:6716, 2012, ISSN: 2070-1721
5. ITU-T Rec. P.86 - Perceptual objective listening quality assessment (2014)
6. ISO/IEC 13818-1: Information technology - Generic coding of moving pictures and associated audio information: Systems, 2000.
7. ISO/IEC 13818-7: Information technology - Generic coding of moving pictures and associated audio information - Part 7: Advanced Audio Coding (AAC), 2004.
8. Peter Pocta, John G. Beerends, Subjective and objective assessment of perceived audio quality of current digital audio broadcasting systems and web-casting applications - IEEE Transactions on broadcasting, vol. 61, no. 3, september 2015 – pp. 407 – 415, DOI: 10.1109/TBC.2015.2424373
9. Giancarlo Vercellesi, Martino Zerbini, Andrea Lorenzo Vitali - Objective and subjective evaluation MPEG layer III perceived quality, 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006
10. Giovanni Bucci, Federico Franciosi, Pietro Valocchi, The measurement of audio codec sound quality, IEEE instrumentation and Measurement Technology Conference, Brussels, Belgium, June 4-6, 1996, pp.622-627
11. Marija Salovarda, Ivan Bolkovac, Hrvoje Domitrovic, Comparsion of audio codecs using PEAQ algorithm, URL:

https://www.researchgate.net/publication/267774023_Comparison_of_audio_codecs_using_PEAQ_algorithm

12. Marina Bosi, Richard E. Goldberg, Introduction to Digital Audio Coding and Standards, SPRINGER SCIENCE+BUSINESS MEDIA, LLC, 2003, ISBN:978-1-4613-5022-4
13. Котвицький І. В., Оса А. Об'єктивне та суб'єктивне оцінювання якості музичних сигналів обмежених смугою частот. – Електронна та акустична інженерія, том 1, № 1, 2018, сс. 34 – 38, ISSN 2524-2725
14. Karlheinz Brandenburg, MP3 and AAC explained, AES 17th International Conference on High Quality Audio Coding, Signa, Italy, September, 1999
15. Продеус А., Дидковский В., Дидковская М. Акустическая экспертиза и коррекция коммуникационных каналов. Монография. – LAP LAMBERT Academic Publishing, OmniScriptum GmbH & Co. KG, Saarbrücken, Deutschland, 2017, ISBN: 978-3-330-04591-0
16. A. Prodeus, “Reducing Sensitivity of Segmental Signal-to-Noise Ratio Estimator to Time-Alignment Error” Int. J. Electr. Electron. Sci., vol. 2, no. 2, pp. 31–36, 2015.
URL: <http://www.aascit.org/journal/archive2?journalId=915&paperId=2822>
17. A. Prodeus, “On Some Features of Log-Spectral Distortion as Speech Quality Measure,” Autom. Softw. Dev. Eng. J., vol. 1, 2016.
URL: <https://asdej.xyz/asde-journal-vol-1/>
18. Recommendation ITU-T P.913. Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment. Edition 2.0., 2016.
19. ITU–T Rec. P.830. Telephone transmission quality. Methods for objective and subjective assessment of quality. Subjective performance assessment of telephone-band and wideband digital codecs, 1996.
20. ITU–R Rec. BS.1116–1. Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems, 1997.

- 21.ITU–R Rec. BS.1284–1. General methods for the subjective assessment of sound quality, 2003.
- 22.ITU-R Rec. BS.1387-1 Method for objective measurements of perceived audio quality (1998-2001)
- 23.ITU-T Rec. P.800 –Methods for subjective determination of transmission quality (1996)
- 24.Audio Quality Measurement Primer. Application Note. AN9789 Rev 0.00., February,1998,URL:
<https://www.intersil.com/content/dam/intersil/documents/an97/an9789.pdf>
- 25.Klaus Schiffner, Psychoacoustic Audio Quality Measurements Using R&S UPV Audio Analyzer. Application Note, 2009, URL:
https://www.rohde-schwarz.com/ua/applications/application-note_56280-15608.html
- 26.T.Thiede, W.Treurniet, R.Bitto, C.Schmidmer, T.Sporer, J.Beerends, C. Colomes, M.Keyhl, G.Stoll, K.Brandenburg, B.Feiten, PEAQ—the ITU standard for objective measurement of perceived audio quality, J. Audio Eng. Soc., vol.:48, 1/2, pp.3–29, 2000.
- 27.DermotCampbell, EdwardJones, MartinGlavin, Audio quality assessment techniques—A review, and recent developments, Signal Processing 89 pp.1489–1500, 2009, ISSN: 0165-1684
- 28.Kabal, P. An Examination and Interpretation of ITU-R BS.1387: Perceptual Evaluation of Audio Quality. TSP Lab Technical Report, Dept. ECE, McGill Univ. May, 2002.
- 29.Dinu Câmpeanu and Andrei Câmpeanu, PEAQ – an objective method to assess the perceptual quality of audio compressed files URL:
https://www.researchgate.net/publication/228944316_PEAQ-An_Objective_Method_To_Assess_The_Perceptual_Quality_of_Audio_Compressed_Files